In the three sequences deduced there are still five glutamic acid residues so that one must be common to two sequences. This can only be the *C*-terminal residue of sequence 2 and the *N*-terminal residue of sequence 3, which are therefore joined together, and the structure in fraction *A* becomes Ser.Leu.Tyr.Glu.Leu.Glu.Asp.Tyr.$CySO_3H$. Both the 'Leu' residues in this sequence are leucine and not isoleucine since peptides $A_{1\iota 7}$ (Ser.Leu) and $A_{1\delta 6}$ (Glu.Leu.Glu) contain no isoleucine. The presence of peptide A_{1a17} ([Glu, Tyr, Leu]), which was different from A_{1a20} (Leu.Tyr.Glu), can readily be understood from this sequence. A number of structures are possible, such as Tyr.Glu.Leu or Leu.Tyr.Glu.Leu.Glu.

All the peptides encountered in this work are now accounted for in terms of the four sequences:

(1) Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala.
(2) Ser.Leu.Tyr.Glu.Leu.Glu.Asp.Tyr.$CySO_3H$.
(3) Ser.Val.$CySO_3H$.
(4) $CySO_3H$.Asp.

All the amino-acids known to be present in fraction *A* are present in these sequences and there is one too many cysteic acid residues. The $CySO_3H$.Asp must therefore be joined to the *C*-terminal end of sequence 2 or 3, but it is not possible to decide conclusively which.

In conclusion, the various peptides identified are listed in Table 14 with the sequences deduced from them.

SUMMARY

1. Fraction *A* of oxidized insulin has been subjected to partial hydrolysis with acid and the resulting peptides fractionated by charcoal adsorption, ionophoresis and paper chromatography.

2. It is concluded that the following sequences are present in this fraction:

Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala,
Ser.Leu.Tyr.Glu.Leu.Glu.Asp.Tyr.$CySO_3H$,
Ser.Val.$CySO_3H$ and $CySO_3H$.Asp.

One of us (E. O. P. T.) wishes to thank the Commonwealth Government of Australia for a Commonwealth Scientific and Industrial Research Organization studentship. Part of the work described in this paper was carried out during the tenure by F.S. of a Beit Memorial Fellowship, which is gratefully acknowledged.

REFERENCES

Andrews, S. & Schmidt, C. L. A. (1927). *J. biol. Chem.* **73**, 651.
Biserte, G. & Osteaux, R. (1951). *Bull. Soc. Chim. biol., Paris*, **33**, 50.
Blackburn, S. & Lowther, A. G. (1951). *Biochem. J.* **48**, 126.
Chibnall, A. C. & Rees, M. W. (1951). *Biochem. J.* **48**, xlvii.
Consden, R. & Gordon, A. H. (1950). *Biochem. J.* **46**, 8.
Consden, R., Gordon, A. H. & Martin, A. J. P. (1946). *Biochem. J.* **40**, 33.
Consden, R., Gordon, A. H. & Martin, A. J. P. (1947). *Biochem. J.* **41**, 590.
Consden, R., Gordon, A. H. & Martin, A. J. P. (1948). *Biochem. J.* **42**, 443.
Consden, R., Gordon, A. H. & Martin, A. J. P. (1949). *Biochem. J.* **44**, 548.
Durrum, E. L. (1950). *J. Amer. chem. Soc.* **72**, 2943.
Gutfreund, H. & Ogston, A. G. (1949). *Biochem. J.* **44**, 163.
Kritchevsky, T. H. & Tiselius, A. (1951). *Science*, **114**, 299,
Monnier, R. & Penasse, L. (1950). *C.R. Acad. Sci., Paris*, **230**, 1176.
Sanger, F. (1949*a*). *Biochem. J.* **44**, 126.
Sanger, F. (1949*b*). *Biochem. J.* **45**, 563.
Sanger, F. & Tuppy, H. (1951*a*). *Biochem. J.* **49**, 463.
Sanger, F. & Tuppy, H. (1951*b*). *Biochem. J.* **49**, 481.
Tristram, G. R. (1949). *Advanc. Protein Chem.* **5**, 83.
Woolley, D. W. (1948). *Fed. Proc.* **7**, 200.

The Amino-acid Sequence in the Glycyl Chain of Insulin

2. THE INVESTIGATION OF PEPTIDES FROM ENZYMIC HYDROLYSATES

BY F. SANGER* AND E. O. P. THOMPSON†
Biochemical Laboratory, University of Cambridge

(*Received* 16 *July* 1952)

In the previous paper (Sanger & Thompson, 1953*a*) four sequences were shown to be present in fraction *A* of oxidized insulin. It was found, as in the case of fraction *B* (Sanger & Tuppy, 1951*a*), that the complete sequence could not be derived from the small peptides produced on acid hydrolysis, due largely to the great lability of the bonds involving the amino groups of the serine residues. It was therefore necessary to investigate the action of proteolytic enzymes, and the present paper describes the results obtained with peptic and chymotryptic hydrolysates. Trypsin was found to be without action on fraction *A*.

* Member of the scientific staff of the Medical Research Council.
† Present address: Wool Textile Research Laboratories, 343 Royal Parade, Parkville, Victoria, Australia.

Although, in fact, it was only necessary to consider the structure of one of the peptides (A_{p5}) to work out the sequence of fraction *A*, the hydrolysates were investigated in considerable detail, since confirmation was desired and since the results throw some light on the specificity of the enzymes concerned. Also a number of the peptides from these hydrolysates were required to determine the position of the amide groups (Sanger & Thompson, 1953*b*). The action of carboxypeptidase on fractions *A* and *B*, insulin, and acetyl insulin has also been studied since this was used to confirm the presence of asparagine as the *C*-terminal residue of the glycyl chain.

The peptides with reference numbers starting A_1 or A_2 refer to those described in the previous paper.

METHODS

Fraction *A* was prepared free of TyrX as previously described (Sanger & Thompson, 1953*a*). The pepsin used was crystalline pig pepsin (Armour Laboratories). The chymotrypsin was prepared by Mr J. Lightbown of the National Institute for Medical Research. The carboxypeptidase was a commercial preparation obtained from Armour Laboratories. Acetylinsulin was prepared by the method of Hughes (1947).

Experiment A_p (pepsin)

In a typical experiment 50 mg. fraction *A* were dissolved in 5 ml. 0·1 N-acetic acid and 2 mg. pepsin added. The mixture was incubated at 37° for 48 hr. After boiling to inactivate the enzyme, the solution was centrifuged to remove denatured pepsin and evaporated to dryness *in vacuo*. In some experiments the digestion was carried out in 0·01 N-HCl, but no difference in the peptides obtained could be detected.

No preliminary fractionation was necessary and samples of the hydrolysate were investigated directly on paper chromatograms. Where Whatman no. 4 filter paper was used an amount equivalent to 5 mg. of the original fraction *A* was applied, and for Whatman no. 3 paper 10–12·5 mg.

The peptides were investigated essentially by the methods described in the previous paper.

Experiment A_c (chymotrypsin)

Fraction *A* (50 mg.) was dissolved in 5 ml. water, 2·5 mg. chymotrypsin were added and the pH adjusted to 7·5 with dilute NH_3. The mixture was incubated at 37° for 24 hr. It was boiled to inactivate the enzyme and evaporated to dryness *in vacuo*. Before application to a chromatogram, the residue was taken up in a small volume of water, and centrifuged to remove insoluble material.

Experiments with carboxypeptidase

In preliminary experiments it was evident that the carboxypeptidase preparation contained traces of chymotrypsin, which were splitting fraction *A* in the chain. In order to avoid this the enzyme was first incubated with di*iso*propyl fluorophosphonate, which is a specific inhibitor for chymotrypsin (Jansen, Nutting, Jang & Balls, 1949). The action of carboxypeptidase was not affected by this reagent. In order to remove the long-chain material, which frequently interfered with the paper chromatograms, the molecular sieve procedure of Thompson & Partridge (1952) was used.

The di*iso*propyl fluorophosphonate was kept as a 0·1 M solution in dry *iso*propanol. When required for use it was brought to 0·001 M by the addition of water. Of a suspension of carboxypeptidase in water (10–15 mg./ml.) 0·05 ml. was run into 1·45 ml. of this inhibitor solution, and, after adjusting the pH to 7·5, the mixture was incubated at 37° for 30 min. The solution was then added to the substrate (fraction *A*, 10 mg.; fraction *B*, 15 mg.; insulin and acetyl insulin, 25 mg.). Samples (0·3 ml.) were withdrawn at various times with a vacuum pipette. These were heated in a boiling-water bath for 5 min. to inactivate the enzyme. To each were added approximately 50 mg. sulphonated polystyrene resin, which was in bead form (15–30 mesh/in.) and had been regenerated before use by successive washings with 2 N-HCl, water, 2 N-NaOH, water, 2 N-HCl and water (Thompson & Partridge, 1952). The mixture was shaken for 1 hr., the supernatant solution decanted and the beads washed twice with distilled water. The amino-acids were then displaced from the resin with 5 N-NH_3 (0·2 ml.) and investigated on paper chromatograms.

RESULTS

Experiment A_p (pepsin)

Two chromatograms of the peptic hydrolysate are shown in Figs. 1 and 2. When each solvent was run only one length of the paper (Fig. 1) spots 3–9 were not sufficiently resolved to give clear results. It was therefore necessary to clamp a second folded sheet of paper on to the bottom edge during each run so that the chromatograms were developed two lengths in each direction (Fig. 2). The combined results are shown in Table 1.

Partial hydrolysis of peptides A_{p5}, A_{p14}, A_{p15}. Eluates of spot A_{p5} from two chromatograms on Whatman no. 3 paper, were hydrolysed with 12 N-HCl at 37° for 3 days and the hydrolysate was chromatographed on Whatman no. 4 paper. The results are shown in Fig. 3 and Table 2.

Eluates of spots A_{p14} and A_{p15} were treated with 1:2:4-fluorodinitrobenzene and the dinitrophenyl (DNP) derivatives hydrolysed 3 days with 12 N-HCl at 37°. The hydrolysates were extracted with ether, and in both cases bis-DNP-tyrosine was identified in the extract by using reversed phase paper chromatography. The aqueous solutions were chromatographed on Whatman no. 4 paper. Peptide A_{p14} gave on partial hydrolysis Glu.Leu ($A_{1\varepsilon 8}$) Leu.Glu ($A_{1\varepsilon 7}$) and Glu.Leu.Glu ($A_{1\delta 6}$) in addition to the free amino-acids and consequently had the structure Tyr.Glu.Leu.Glu. Peptide A_{p15} gave only free amino-acids and Glu.Leu ($A_{1\varepsilon 8}$) on partial hydrolysis and must therefore have been Tyr.Glu.Leu.

Table 1. *Peptides from peptic hydrolysate (A_p) of fraction A of oxidized insulin*

Spot no. (Figs. 1 and 2)	Amino-acids identified	Strength	DNP amino-acids identified	Structure*
1	$CySO_3H$	× × ×		
	Asp	× × × ×	Glu	Glu.[$CySO_3H$, Asp, Tyr]
	Glu	× × ×		
	Tyr	× × ×		
2	$CySO_3H$	×		
	Asp	× ×	Glu	Glu.[$CySO_3H$, Asp, Tyr]
	Glu	×		
	Tyr	×		
3	$CySO_3H$	× × ×		
	Asp	× × × ×	Asp	Asp.[$CySO_3H$, Asp, Tyr]
	Tyr	× × ×		
4	$CySO_3H$	×		
	Asp	× ×	Asp	Asp.[$CySO_3H$, Asp, Tyr]
	Tyr	×		
5	$CySO_3H$	× × × × ×		
	Glu	× × × × ×		
	Ser	× × × × ×		
	Gly	× × × ×	Gly	Gly.[$CySO_3H$, Glu, Ser, Ala, Val, Leu]
	Ala	× × × ×		
	Val	× × × × ×		
	Leu	× × × × ×		
6	$CySO_3H$	× ×		
	Asp	× ×		
	Glu	× ×	Leu†	Leu.[$CySO_3H$, Asp, Glu, Tyr]
	Tyr	× ×		
	Leu	× ×		
7	$CySO_3H$	×		
	Asp	× ×		
	Glu	×		[$CySO_3H$, Asp, Glu, Tyr, Leu]
	Tyr	×		
	Leu	×		
8	Glu	×		Glutamic acid
9	$CySO_3H$	× ×		
	Glu	× ×		
	Ser	× ×	Gly	Gly.[$CySO_3H$, Glu, Ser, Ala, Val, Leu]
	Gly	× ×		
	Ala	× ×		
	Val	× ×		
	Leu	× ×		
10	Glu	× ×		[Glu, Tyr]
	Tyr	× ×		
11	Glu	× ×		[Glu, Tyr]
	Tyr	× ×		
12	Tyr	×		Tyrosine
13	Glu	× ×		Leu.Glu
	Leu	×		
14	Glu	× × × ×		
	Tyr	× × ×	Tyr†	Tyr.[Glu, Leu]
	Leu	× × ×		
15	Glu	× × ×		
	Tyr	× × ×	Tyr†	Tyr.[Glu, Leu]
	Leu	× × ×		
16	Glu	×		
	Gly	×		Gly.Ileu.Val.Glu ($A_{1/3}$)
	Val	×		
	Leu	×		

* Where amino-acids are enclosed in square brackets the sequence has not been directly established.
† Identified by reversed-phase paper chromatography.
Spot 5. Yellow colour with ninhydrin.
Spot 9. Red-yellow colour with ninhydrin.
Spot 13. The aqueous solution of a hydrolysate of the DNP derivative contained Glu (× ×) Leu (×) suggesting it is Leu.Glu.

Experiment A_c (chymotrypsin)

Fig. 4 shows a chromatogram of the chymotryptic hydrolysate. Since there was considerable overlapping of some of the spots, each one was purified by ionophoresis on paper in 0·05 M-ammonium acetate using the technique of Durrum (1950). With the exception of spot (6+7), only one main component was found in each case and the amino-acids given on hydrolysis are recorded in Table 3. In the case of spot (6+7) two bands were visible after ionophoresis when the paper was examined in ultraviolet light. One (A_{p6}) was neutral and the other (A_{p7}) was acidic. A_{p6} gave a yellow colour with ninhydrin, whereas no colour was obtained with A_{p7}. The composition of the two is shown in Table 3.

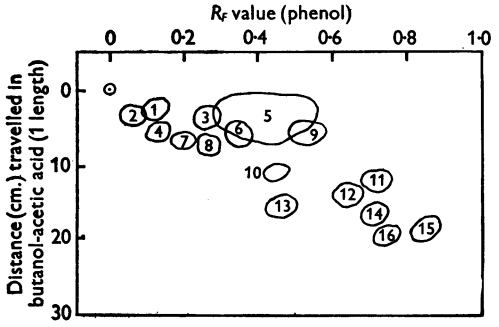

Fig. 1. Chromatogram of peptic hydrolysate of fraction A of oxidized insulin (Exp. A_p) in which both solvents were run one length of paper (see Table 1).

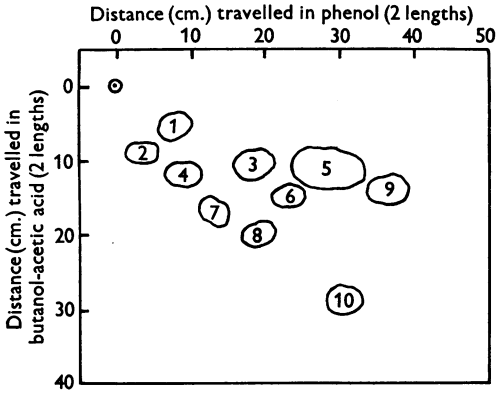

Fig. 2. Chromatogram of peptic hydrolysate of fraction A of oxidized insulin (Exp. A_p) in which both solvents were run two lengths of paper (see Table 1).

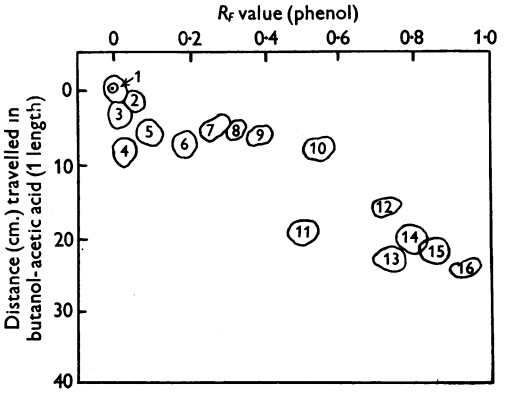

Fig. 3. Chromatogram of partial acid hydrolysate of peptide A_{p5} (see Table 2).

In later experiments, when further samples of these peptides were required, it was more convenient to run a large-scale ionophoresis first, followed by two-dimensional chromatography of the bands obtained. Fig. 5 is a diagram of such an ionophoresis, showing the relative positions of the different peptides.

Table 2. *Peptides from partial hydrolysate of peptide A_{p5}*

Spot no. (Fig. 3)	Strength	Structure	Identical with
1	×	$CySO_3H.CySO_3H.Ala$ + $Glu.CySO_3H.CySO_3H.Ala$	$A_{2\delta1}$ + $A_{2\gamma1}$
2	×	Cysteic acid	—
3	×	$Glu\ CySO_3H$	$A_{1\alpha1}$
4	×	Glu Glu	$A_{1\beta4}$
5	× ×	$CySO_3H.Ala$	$A_{1\gamma3}$
6	×	Glutamic acid	—
7	× × ×	$Ser.Val.CySO_3H$	$A_{1\zeta2}$
8	× ×	Serine	—
9	× ×	Glycine	—
10	× ×	Alanine	—
11	× ×	Gly.Ileu.Val.Glu.Glu. + Ileu Val Glu.Glu	$A_{1\alpha15}$ $A_{1\alpha16}$
12	×	Ser.Val	$A_{1\iota6}$
13	×	Ileu.Val.Glu	$A_{1\eta4}$
14	× ×	Ser Leu	$A_{1\iota7}$
15	×	Leucine + isoleucine	—
16	?	Ileu.Val	$A_{1\iota10}$

Table 3. *Peptides from chymotryptic hydrolysate (A_c) of fraction A of oxidized insulin*

Spot no. (Fig. 4)	Amino-acids formed after hydrolysis: $CySO_3H$	Asp	Glu	Ser	Gly	Ala	Tyr	Val	Leu
1	× × ×	× × ×	—	—	—	—	—	—	—
2	× × ×	—	× × ×	× × ×	× × ×	× × ×	—	× × ×	× ×
3	× ×	× × ×	× × × ×	—	—	—	× ×	—	× ×
4	× × × ×	—	× × × ×	× × × ×	× × × ×	× × × ×	× × ×	× × × ×	× × × ×
5	—	× ×	× × ×	—	—	—	× ×	—	× ×
6	—	—	—	× ×	—	—	× × ×	—	× × ×
7	—	× ×	× × ×	—	—	—	× ×	—	× ×

Spot 1. Yellow colour with ninhydrin, turning to purple on prolonged heating. Hydrolysis of the DNP derivative gives DNP-cysteic acid and aspartic acids. Its structure is therefore $CySO_3H$.Asp.
Spot 2. Yellow colour with ninhydrin. Contained isoleucine but no leucine.
Spot 3. Hydrolysis of the DNP derivative gave DNP-glutamic acid.
Spot 4. Yellow colour with ninhydrin. Contained both leucine and isoleucine.

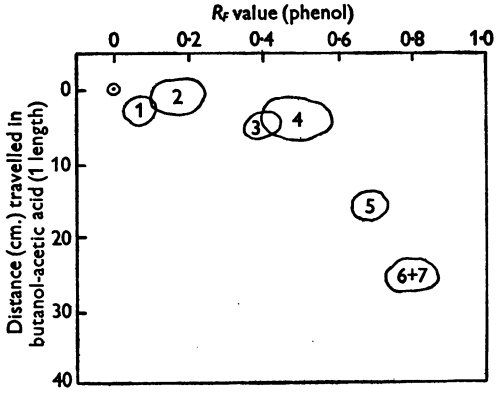

Fig. 4. Chromatogram of chymotryptic hydrolysate of fraction *A* of oxidized insulin (Exp. A_c) (see Table 3).

Table 4. *Action of carboxypeptidase on insulin derivatives*

Substrate	Time of incubation at 37°	Free amino-acids liberated	Strength
Fraction *B*	30 min.	Ala	× × × × ×
Fraction *A*	1 hr.	Asp (-NH_3)	×
		Ala	?
Insulin	1·5 hr.	Ala	× × × × ×
		Asp (-NH_3)	× ×
		Asp	?
	8 hr.	Ala	× × × × ×
		Asp (-NH_3)	× × ×
		Asp	×
Acetylinsulin	4 hr.	Ala	× × × × ×
		Asp (-NH_3)	× × × ×
		Asp	×

Carboxypeptidase experiments

The results obtained with carboxypeptidase are summarized in Table 4. Free alanine was liberated very rapidly from fraction *B* and insulin, as has been shown by Lens (1949). No evidence could be found of free lysine which is the residue adjacent to alanine (Sanger & Tuppy, 1951*b*). The first amino-acids liberated from fraction *A* were asparagine and a trace of alanine, which is probably due to contamination with fraction *B*. On prolonged incubation traces of glutamic acid, serine, tyrosine and leucine were liberated. Better yields of asparagine were obtained from insulin. This effect may be due to inhibition of the enzyme by the —SO_3H group of the adjacent cysteic acid residue in fraction *A*. An approximate estimation by the spot-dilution technique of Polson, Mosley & Wyckoff (1947) indicated

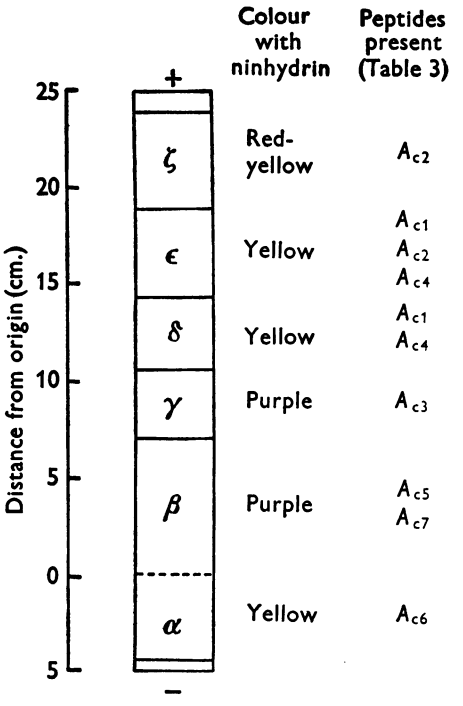

Fig. 5. Ionophoresis of chymotryptic hydrolysate of fraction *A* of oxidized insulin in 0·05 M-ammonium acetate (see Table 3).

that after 8 hr. incubation the yield of asparagine from insulin was 25–30% of the theoretical value, assuming two *C*-terminal asparagine residues per molecule of mol.wt. 12000.

Acetylinsulin was investigated in the hope that the lysine residue from the phenylalanyl chain could be split off as ε-acetyllysine. This expectation was not realized, but a much better yield of asparagine was obtained. After a reaction period of 4 hr. the yields of both alanine and asparagine were approximately 90% of the theoretical value. Aspartic acid was also obtained in a yield of less than 10%. It is probably a secondary product derived from asparagine, although control experiments in which asparagine was incubated with carboxypeptidase showed that the breakdown was very small and would hardly account for the amounts of aspartic acid formed in the experiments described above.

DISCUSSION

Amino-acid sequence of fraction A

In the previous paper four sequences were deduced as being present in fraction *A*. These were:

(1) Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala (*N*-terminal sequence).
(2) Ser.Leu.Tyr.Glu.Leu.Glu.Asp.Tyr.$CySO_3H$.
(3) Ser.Val.$CySO_3H$.
(4) $CySO_3H$.Asp.

The peptide present in the highest concentration in the peptic hydrolysate was peptide A_{p5}. It contained all the amino-acids of fraction *A* except aspartic acid and tyrosine, and had a glycine *N*-terminal residue. On partial hydrolysis with acid it gave the peptides listed in Table 2. Of these, all but peptides 7 (Ser.Val.$CySO_3H$), 12 (Ser.Val) and 14 (Ser.Leu) are derived from sequence 1. Since peptide A_{p5} gives Ser.Leu on hydrolysis but contains no tyrosine, this leucine residue must be present as the *C*-terminal residue. It contains therefore sequence 1 as its *N*-terminal sequence, Ser.Leu as its *C*-terminal sequence and Ser.Val.$CySO_3H$ which must be within the chain. Neither $CySO_3H$.Asp nor the remainder of sequence 2 can be present in A_{p5}, so that its structure must be Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala.Ser.Val.$CySO_3H$.Ser.Leu. Having established this sequence containing three cysteic acid residues there is only one such residue to which $CySO_3H$.Asp can be attached and the structure of fraction *A* must be Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala.Ser.Val.$CySO_3H$.Ser.Leu.Tyr.Glu.Leu.Glu.Asp.Tyr.$CySO_3H$.Asp.

For simplicity in deducing this sequence, considerable use has been made of the figures given in the previous paper (Sanger & Thompson, 1953*a*, Table 13) for the amino-acid composition of fraction *A*. It could, however, have been deduced as the simplest structure which would fit all the experimental results obtained, and, in fact, was originally worked out in this way. For instance, it could be shown that all the tyrosine-containing peptides could not be explained on the basis of one sequence, but could be explained on the basis of two sequences. This is fairly conclusive evidence that there are two tyrosine residues. If there were three it is unlikely that all the peptides could be fitted into two sequences. While the structure deduced is the simplest one that will fit the experimental results, a few more complicated ones are possible but are ruled out by the analytical figures. For instance, if there were an extra serine residue between positions 8 and 9 (Table 5) this might not have been detected, but the analytical results show that there are only two such residues in the chain. It is seen that an exact amino-acid analysis of the chain may in this way be worked out from the nature of the peptides without carrying out any quantitative estimation. It is interesting that the very approximate analysis of fraction *A* (Table 13, Sanger & Thompson, 1953*a*) which was done by paper chromatography is in fact correct except for the leucine value which should be 2 and not 3.

The fact that it is possible to deduce this single sequence proves that fraction *A* is an essentially homogeneous peptide and therefore that there is only one type of glycyl chain in insulin.

As was the case with the phenylalanyl chain, there is no obvious periodicity of the amino-acids in this chain. In fraction *B* it was noted that the aromatic and less polar residues were grouped together. There is no such arrangement in fraction *A*, so that this cannot be regarded as a general property of proteins. In fraction *B* there were three dipeptide sequences which occurred twice in the chain, and this was considerably more than would be expected on a random arrangement. No dipeptide sequences occur twice in the glycyl chain, although there are now six which occur twice in the whole insulin. These are Val.Glu (*A* 3–4, *B* 12–13), Val.$CySO_3H$ (*A* 10–11, *B* 18–19), Leu.Tyr (*A* 13–14, *B* 15–16), His.Leu (*B* 5–6, *B* 10–11), Gly.$CySO_3H$ (*B* 7–8, *B* 19–20) and Leu.Val (*B* 11–12, *B* 17–18). The sequence CyS.Gly, which occurs in glutathione and twice in the phenylalanyl chain, is not found in the glycyl chain. It would thus seem that no general conclusions can be drawn from these results concerning the general principles which govern the arrangement of amino-acid residues in protein chains. In fact, it would seem more probable that there are no such principles, but that each protein has its own unique arrangement; an arrangement which endows it with its particular properties and specificities and fits it for the function that it performs in nature.

Table 5. *Peptides obtained by enzymic hydrolysis of fraction A of oxidized insulin*

Sequences deduced from lower peptides (Sanger & Thompson, 1953*a*)	Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala Ser.Val.$CySO_3H$ Ser.Leu.Tyr.Glu.Leu.Glu.Asp.Tyr.$CySO_3H$ $CySO_3H$.Asp
Peptides from peptic hydrolysate	Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala.Ser.Val.$CySO_3H$.Ser.Leu (A_{p5}) Leu.Glu.Asp.Tyr.$CySO_3H$.Asp (A_{p6}, A_{p7}) Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala.Ser.Val (A_{p9}) Tyr.Glu (A_{p10}, A_{p11}) Glu.Asp.Tyr.$CySO_3H$.Asp (A_{p1}, A_{p2}) Gly.Ileu.Val.Glu (A_{p16}) Tyr.Glu.Leu (A_{p15}) Asp.Tyr.$CySO_3H$.Asp (A_{p3}, A_{p4}) Tyr.Glu.Leu.Glu (A_{p14}) Leu.Glu (A_{p13})
Peptides from chymotryptic hydrolysate	Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala.Ser.Val.$CySO_3H$.Ser.Leu.Tyr (A_{c4}) $CySO_3H$.Asp (A_{c1}) Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala.Ser.Val.$CySO_3H$ (A_{c2}) Glu.Leu.Glu.Asp.Tyr.$CySO_3H$.Asp (A_{c3}) Ser.Leu.Tyr (A_{c6}) Glu.Leu.Glu.Asp.Tyr (A_{c5}, A_{c7})
Structure of fraction *A*	Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala.Ser.Val.$CySO_3H$.Ser.Leu.Tyr.Glu.Leu.Glu.Asp.Tyr.$CySO_3H$.Asp 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Bonds split by pepsin Bonds split by chymotrypsin	↑, major sites of action of enzymes; †, other bonds split by enzymes.

Structure of the peptides

Having established the sequence of residues in fraction *A* it is now possible to consider the peptides identified in the various hydrolysates. Their structures are given in Table 5. The fact that these can all be determined in this way provides confirmatory evidence for the main sequence.

Since pepsin splits bonds involving the amino groups of tyrosine residues in synthetic peptides (Fruton & Bergmann, 1939), it was expected that splitting of fraction *A* would occur at position 18–19 to give Tyr.$CySO_3H$.Asp. Although peptides containing these three amino-acids were present (A_{p3}, A_{p4}), they had aspartic acid and not tyrosine as the *N*-terminal residues. Splitting by pepsin therefore occurred at bond 17–18 (Glu.Asp) and no evidence could be obtained for splitting of bond 18–19. Further considerations on the specificity of pepsin will be presented in a later paper (Sanger & Thompson, 1953*b*), when the positions of the amide groups have been established.

The presence of duplicate spots of the peptides Asp.Tyr.$CySO_3H$.Asp (A_{p3}, A_{p4}), Glu.Asp.Tyr.-$CySO_3H$.Asp (A_{p1}, A_{p2}) and Leu.Asp.Tyr.-$CySO_3H$.Asp (A_{p6}, A_{p7}) is probably due to deamidation. Amide groups are present on the aspartic acid residues (Sanger & Thompson, 1953*b*) and some hydrolysis probably occurred during the digestion with pepsin in acid solution (cf. Melville, 1935). Similarly, peptides A_{p10} and A_{p11} (Tyr.Glu) differ by an amide group.

Partial acid hydrolysis of peptide A_{p9} on a rather small amount of material suggested the structure shown in Table 5, since it gave the peptides from the *N*-terminal sequence and Ser.Val but no Ser.Val.-$CySO_3H$ or Ser.Leu.

In the chymotryptic hydrolysate two peptides (A_{c5}, A_{c7}) containing the same amino-acids (Asp, Glu, Tyr, Leu) are present. A_{c7} moves faster on the chromatograms than A_{c5} and gives no colour with ninhydrin, suggesting that it has no free α-amino group. Both are derived from the sequence Glu.Leu.Glu.Asp.Tyr and evidence to be presented later (Sanger & Thompson, 1953*b*) indicates that this *N*-terminal glutamic acid residue is present in the form of an amide. Glutamine residues containing a free α-amino group are particularly labile and readily cyclize to give the corresponding pyrollidonecarboxylic acid derivative (Melville, 1935). It seems probable that peptide A_{c7} is such a derivative formed from A_{c5} during the digestion or subsequent operations.

Peptide A_{c6} can only be Ser.Leu.Tyr identical with A_{1a19}, indicating that splitting with chymotrypsin has occurred at bond 11–12 ($CySO_3H$.Ser). This is confirmed by the presence of peptide A_{c2} which contains isoleucine and the other amino-acids from the *N*-terminal sequence but no leucine. Since the splitting of this bond was rather unexpected, control experiments were carried out to show that the effect was due to chymotrypsin and not to the experimental conditions used. An end-group determination on the fraction *A* used showed that there were no free amino groups of serine present, glycine being the only *N*-terminal residue. Samples of fraction *A* were then treated in the usual manner with chymotrypsin which had been previously incubated at 37° for 30 min. with a 0·001 M solution of the specific inhibitor di*iso*propylfluorophosphonate. Chromatography of the evaporated solutions showed no spot corresponding to peptide A_{c2}. Only one large area (R_F in phenol = 0·25) was obtained on the chromatograms which gave a yellow colour with ninhydrin and on hydrolysis gave all the amino-acids of fraction *A*. It was obviously unchanged fraction *A*. It is therefore concluded that the splitting of this bond is due to the action of chymotrypsin or possibly of another enzyme that is very closely associated with it and which is inhibited by di*iso*propyl fluorophosphonate.

Further evidence for this splitting was provided in an experiment in which peptide A_{p5} was treated with chymotrypsin. Ser.Leu was identified as the only neutral component after ionophoresis in 0·05 M-ammonium acetate. The bands moving towards the anode were separated by chromatography and shown to be similar in R_F values and composition to A_{c2} and A_{p5} (unchanged peptide).

The C-terminal residue of fraction A

Peptide A_{c1} contains the same amino-acid sequence as peptide A_{1a2} ($CySO_3H$.Asp). However, it was found to behave differently on ionophoresis in 0·05 M-ammonium acetate. Thus after ionophoresis for 6 hr. with a potential of 220 V., A_{1a2} had moved 8·4 cm., whereas A_{c1} had only moved 5·1 cm. This clearly shows that A_{c1} is an amide. Its structure is therefore $CySO_3H$.Asp(-NH_2) and the *C*-terminal residue of fraction *A* is asparagine. Although there is now considerable confirmatory evidence for this (Chibnall & Rees, 1952; Harris, 1952), it had been suggested at the time this sequence was first being worked out, that glycine was present as a *C*-terminal residue in insulin, presumably on the glycyl chains (Fromageot, Jutisz, Meyer & Penasse, 1950; Chibnall & Rees, 1951). As a result of more recent work Chibnall & Rees (private communication) have now withdrawn this claim.

The analytical results on fraction *A* indicated that there was only one glycine residue and this appeared to be the *N*-terminal residue. This was supported by the results of Butler, Phillips, Stephen & Creeth (1950), which we have been able to confirm. They showed that no glycine was present in

a hydrolysate of the DNP derivative of fraction *A*, showing that the only glycine residues are *N*-terminal.

The results with carboxypeptidase support the conclusion that asparagine is the *C*-terminal residue of the glycyl chain. In the case of acetylinsulin the liberation was approximately quantitative, assuming that it is present in both the glycyl chains, if one assumes that the molecular weight is 12000.

SUMMARY

1. Fraction *A* of oxidized insulin was hydrolysed with pepsin and chymotrypsin. The resulting peptides were fractionated and their structure investigated.

2. It is concluded that the amino-acid sequence in the glycyl chain of insulin is Gly.Ileu.Val.Glu.Glu.CyS.CyS.Ala.Ser.Val.CyS.Ser.Leu.Tyr.Glu.Leu.Glu.Asp.Tyr.CyS.Asp.

We wish to express our thanks to Mr J. Lightbown for the chymotrypsin and to Dr B. C. Saunders for the di*iso*propyl fluorophosphonate.

One of us (E. O. P. T.) wishes to thank the Commonwealth Government of Australia for a Commonwealth Scientific and Industrial Research Organization studentship.

Part of the work described in this paper was carried out during the tenure by F. S. of a Beit Memorial Fellowship.

REFERENCES

Butler, J. A. V., Phillips, D. M. P., Stephen, J. M. L. & Creeth, J. M. (1950). *Biochem. J.* **46**, 74.

Chibnall, A. C. & Rees, M. W. (1951). *Biochem. J.* **48**, xlvii.

Chibnall, A. C. & Rees, M. W. (1952). *Biochem. J.* **52**, iii.

Durrum, E. L. (1950). *J. Amer. chem. Soc.* **72**, 2943.

Fromageot, C., Jutisz, M., Meyer, D. & Penasse, L. (1950). *Biochim. biophys. Acta*, **6**, 283.

Fruton, J. S. & Bergmann, M. (1939). *J. biol. Chem.* **127**, 627.

Harris, J. I. (1952). *J. Amer. chem. Soc.* **74**, 2944.

Hughes, W. R. L. (1947). Cited by Herriot, R. M. in *Advanc. Protein Chem.* **3**, 169.

Jansen, E. F., Nutting, M. D., Jang, R. & Balls, A. K. (1949). *J. biol. Chem.* **179**, 189.

Lens, J. (1949). *Biochim. biophys. Acta*, **3**, 367.

Melville, J. (1935). *Biochem. J.* **29**, 179.

Polson, A., Mosley, Y. M. & Wyckoff, R. W. G. (1947). *Science*, **105**, 603.

Sanger, F. & Thompson, E. O. P. (1953*a*). *Biochem. J.* **53**, 353.

Sanger, F. & Thompson, E. O. P. (1953*b*). (Unpublished.)

Sanger, F. & Tuppy, H. (1951*a*). *Biochem. J.* **49**, 463.

Sanger, F. & Tuppy, H. (1951*b*). *Biochem. J.* **49**, 481.

Thompson, A. R. & Partridge, S. M. (1952). *Nature, Lond.*, **169**, 495.

The Branched-chain Fatty Acids of Mutton Fat

2. THE ISOLATION OF (+)-12-METHYLTETRADECANOIC ACID AND OF 13-METHYLTETRADECANOIC ACID

R. P. HANSEN, F. B. SHORLAND AND N. J. COOKE

Fats Research Laboratory, Department of Scientific and Industrial Research, Wellington, New Zealand

(*Received* 13 *June* 1952)

Following the isolation of branched-chain fatty acids from butterfat (Hansen & Shorland, 1950, 1951, 1952; Hansen, Shorland & Cooke, 1951) and from ox fat (Hansen, Shorland & Cooke, 1952*a*), a detailed examination was made of the lipids extracted from the external fatty tissue of old ewes. The occurrence of (+)-14-methylhexadecanoic acid as a constituent of this fat was recently established (Hansen, Shorland & Cooke, 1952*b*, *c*). In this present paper are reported the isolation from mutton fat of two isomeric C_{15} branched-chain saturated fatty acids, (+)-12-methyltetradecanoic acid and 13-methyltetradecanoic acid.

Although the optically active *ante-iso*-acid, (+)-12-methyltetradecanoic acid, has been isolated from wool grease (Weitkamp, 1945), the *iso* acid, 13-methyltetradecanoic acid, has not formerly been found in natural fats.

EXPERIMENTAL

The long crystal spacings reported in this work were determined by means of a Philips Geiger X-ray spectrometer. Except where otherwise indicated, nickel-filtered copper $K\alpha$ radiation was used. Samples were melted on a glass slide and quickly cooled.

All melting points were determined in closed capillaries and are uncorrected.

Analyses were by Weiler and Strauss, Oxford.

The fat used in this investigation was obtained by mincing and then steam-rendering the external fatty tissues